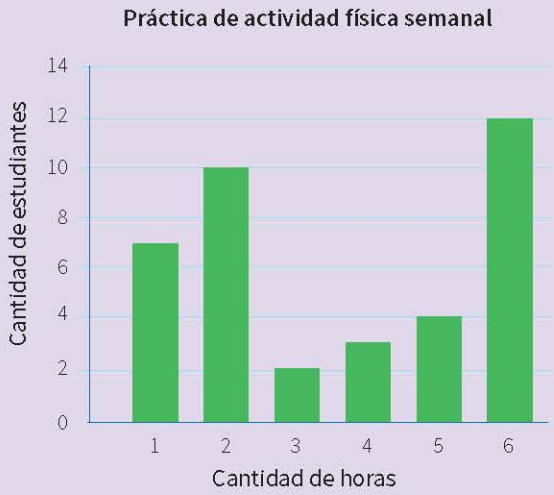
Comprobamos nuestros aprendizajes

Propósito: Representamos las características de una población mediante variables cuantitativas, así como el comportamiento de los datos de una muestra con medidas de tendencia central; además, planteamos conclusiones sobre una población y las justificamos, y corregimos errores si los hubiera.

Situación significativa A

El siguiente gráfico de barras corresponde al número de horas semanales que un grupo de estudiantes de tercero de secundaria dedica a la práctica de actividades físicas.



Calcula la media y la mediana de este conjunto de datos y haz una comparación entre ellas.

Resolución

A partir de la lectura del gráfico de barras, elaboramos la tabla de frecuencias:

Cantidad de horas X_i	f_i	$X_i \cdot f_i$
1	7	$7 \times 1 = 7$
2	10	20
3	2	6
4	3	12
5	4	20
6	12	72
Total	38	137

- Calculamos la media aritmética ($\overline{x}$) que es el promedio de los datos. Para datos no agrupados, se calcula sumando los productos de cada frecuencia absoluta (f_i) por el valor de la variable (X_i) y se divide entre el número de datos:

$$\overline{x} = \frac{\sum_{i=1}^{n} X_i \cdot f_i}{n}$$

Aplicando la fórmula tenemos:

$$\overline{x} = \frac{7(1) + 10(2) + 2(3) + 3(4) + 4(5) + 12(6)}{38} = \frac{137}{38} \approx 3{,}6 \text{ horas.}$$

- Calculamos la mediana (Me), que es el valor numérico que ocupa la ubicación central de la muestra, de modo que el 50 % son menores o iguales que ella y el otro 50 % son mayores o iguales que ella.

Ordenamos los datos de menor a mayor:

1; 1; 1; 1; 1; 1; 1; 2; 2; 2; 2; 2; 2; 2; 2; 2; 2; 3; 3; 4; 4; 4; 5; 5; 5; 5; 6; 6; 6; 6; 6; 6; 6; 6; 6; 6; 6; 6.

Como la cantidad de datos es par ($n = 38$), entonces la muestra se reparte en dos grupos de 19 datos cada uno y la mediana es el promedio de los dos datos centrales.

$\text{Me} = \frac{3 + 4}{2} = 3{,}5$ horas

Finalmente, al comparar encontramos que la media y la mediana se diferencian en solo una décima (0,1).

1. ¿Por qué calculamos la media y la mediana de los datos que se presentan en el gráfico de barras?

2. Considerando que la moda (Mo) es el dato que más se repite, es decir, el que tiene mayor frecuencia absoluta, calcula la moda e interpreta su significado.

Situación significativa B

La siguiente tabla muestra la cantidad de personas, según edad, que se encuentran afiliadas a una asociación del adulto mayor.

Determina la media y la mediana de los datos proporcionados en la tabla.

Edad $[L_i; L_s[$	f_i
[60; 70[	18
[70; 80[	14
[80; 90[	6
[90; 100]	2

Resolución

Con ayuda de la fórmula para datos agrupados, hallaremos la media. Pero antes completamos la tabla de frecuencias con las columnas que vamos a necesitar:

$[L_i; L_s[$	X_i	f_i	F_i	$X_i \cdot f_i$
[60; 70[	65	18	18	1170
[70; 80[	75	14	32	1050
[80; 90[	85	6	38	510
[90; 100]	95	2	40	190
Total		40		2920

- La mediana (Me) para datos agrupados se calcula aplicando la siguiente fórmula:

$$Me = L_i + \left(\frac{\frac{n}{2} - F_{i-1}}{f_i} \right) \cdot A$$

L_i: límite inferior del intervalo de la clase mediana

F_{i-1}: frecuencia absoluta acumulada del intervalo anterior al intervalo de la clase mediana

f_i: frecuencia absoluta del intervalo de la clase mediana

A: amplitud del intervalo

n: número de datos

Identificamos el intervalo de la clase mediana: $\frac{n}{2} = \frac{40}{2} = 20$, observamos que el dato 20 se encuentra contenido en $F_2 = 32$. Por lo tanto, la clase mediana es [70; 80[

Reemplazamos y calculamos: $Me = 70 + \left(\frac{\frac{40}{2} - 18}{14} \right) \cdot 10 = 70 + \frac{20}{14} \approx 71{,}4$ años.

Por tanto, podemos afirmar que, en dicha asociación, el 50 % de las personas afiliadas tienen edades menores o iguales que 71,4 años y el otro 50 % tienen edades iguales o mayores que 71,4 años.

- La media aritmética $(\overline{x}_i)$ para datos agrupados en intervalos es la suma de los productos de las frecuencias absolutas (f_i) por su marca de clase (X_i), dividida entre el número de datos (n).

$$\overline{X} = \frac{\sum_{i=1}^{n} f_i \cdot X_i}{n}$$

$$\overline{X} = \frac{2920}{40} = 73 \text{ años}$$

Por lo tanto, el promedio de la edad de las personas afiliadas a la asociación del adulto mayor es 73 años.

1. ¿Qué nos indica en la tabla de frecuencias la edad y cómo están representados?

2. ¿Qué información nos brinda la columna que corresponde a la marca de clase (X_i)?

3. ¿Por qué es importante calcular la frecuencia absoluta acumulada (F_i)?

4. ¿Por qué es importante calcular la marca de clase (X_i)?

Situación significativa C

La tabla de frecuencias muestra la población de una ciudad en relación con sus edades. A partir de la lectura de la tabla, calcula la moda.

Edad	f_i
[0; 10[	2 756 259
[10; 20[	2 876 709
[20; 30[	2 383 378
[30; 40[	1 921 716
[40; 50[	1 479 675
[50; 60[	999 795
[60; 70[	644 750
[70; 80[	387 911
[80; 90]	152 632
Total	13 602 825

Aprendemos a partir del error

Resolución

- Como estrategia, usamos la fórmula de la moda (Mo) para datos agrupados: $Mo = L_i + \left(\frac{d_1}{d_1 + d_2}\right)A$

  donde:

  $$d_1 = f_{Mo} - f_{Mo-1}$$
  $$d_2 = f_{Mo} - f_{Mo+1}$$

- Además:

  L_i: límite inferior del intervalo modal

  f_{Mo}: frecuencia absoluta del intervalo modal

  f_{Mo-1}: frecuencia absoluta anterior al intervalo modal

  f_{Mo+1}: frecuencia absoluta posterior al intervalo modal

  A: amplitud del intervalo modal

En primer lugar, ubicamos el intervalo modal, que es aquel con la mayor frecuencia absoluta.

$[L_i; L_s[$	f_i
[0; 10[	2 756 259
[10; 20[← Intervalo modal	2 876 709
[20; 30[	2 383 378
[30; 40[	1 921 716
[40; 50[	1 479 675
[50; 60[	999 795
[60; 70[	644 750
[70; 80[	387 911
[80; 90]	152 632
Total	13 602 825

- Identificamos lo necesario para calcular la moda de estos datos:

$L_i = 20$

$f_{Mo} = 2\,876\,709$

$f_{Mo-1} = 2\,756\,259$

$f_{Mo+1} = 2\,383\,378$

$A = 10$

- Con estos datos calculamos d_1 y d_2.

$d_1 = f_{Mo} - f_{Mo-1} = 2\,876\,709 - 2\,756\,259 = 120\,450$

$d_2 = f_{Mo} - f_{Mo+1} = 2\,876\,709 - 2\,383\,378 = 493\,331$

Luego:

$$Mo = L_i + \left(\frac{d_1}{d_1 + d_2}\right)A = 20 + \left(\frac{120\,450}{120\,450 + 493\,331}\right) \cdot 10$$

$$Mo = 20 + \left(\frac{120\,450}{613\,781}\right) \cdot 10 = 20 + 0{,}196 \cdot (10) = 21{,}96$$

Respuesta: Concluimos que la moda de este conjunto de datos es 21,96 años, es decir, aproximadamente 22 años.

1. ¿El procedimiento realizado para la resolución de la situación significativa es correcto? Corrige si fuera incorrecto.

2. Se ha determinado que el intervalo modal es el segundo intervalo [10; 20[. ¿Será correcto que la moda resulte aproximadamente 22 años?, ¿por qué?